Discrete Mathematics 324 (2014) 78–84

Contents lists available at ScienceDirect

Discrete Mathematics

journal homepage: www.elsevier.com/locate/disc

Note

On the locality of codeword symbols in non-linear codes

Michael Forbes [a,1], Sergey Yekhanin [b]

[a] MIT, United States
[b] Microsoft Research, United States

ARTICLE INFO

Article history:
Received 15 March 2013
Received in revised form 12 November 2013
Accepted 24 January 2014
Available online 4 March 2014

Keywords:
Error-correcting codes
Non-linear codes
Locality

ABSTRACT

Coordinate i of an error-correcting code has locality r if its value is determined by some r other coordinates. Recently an optimal trade-off between information locality of linear codes, code distance, and redundancy has been obtained. Furthermore, for linear codes meeting this trade-off, structure theorems were derived. In this work we generalize the trade-off and structure theorems to non-linear codes.

1. Introduction

We say that a certain coordinate of an error-correcting code has locality r if, when erased, the value at this coordinate can be recovered by accessing at most r other coordinates. Motivated by applications to data storage [3] the authors of [2] introduced (r, d)-codes, which are systematic codes that have distance d and thus tolerate up to $d-1$ erasures, but also have the property that any information coordinate has locality r or less. They established that in all *linear* $[n, k, d]_q$ codes with the (r, d)-property

$$n \geq k + \left\lceil \frac{k}{r} \right\rceil + d - 2. \tag{1.1}$$

In what follows we refer to codes that meet (1.1) with equality as *optimal*. A construction of [1] implies that optimal codes exist for all values of parameters.

While locality of data symbols and code distance are the two primary considerations in the design of codes for data storage applications, locality of parity coordinates is also important. Parity locality (in the class of optimal (r, d)-codes) has been considered in [2]. In the natural setting of $r|k$, the lower-bound argument of [2] yields structure theorems for optimal linear codes. These theorems are particularly strong when $d < r + 3$. In that case they imply tight lower bounds for parity locality.

Coding theory knows many examples of problems where non-linear codes improve upon the best available constructions of linear codes, e.g., [6]. While there is currently no evidence that non-linearity facilitates better (r, d)-codes, the novelty of this regime suggests that further study is required. In particular, it is natural to ask whether non-linearity can help reduce redundancy of (r, d)-codes or parity locality of optimal (r, d)-codes. The first question has been addressed in [4,5] where

E-mail addresses: miforbes@mit.edu (M. Forbes), yekhanin@microsoft.com (S. Yekhanin).

[1] Research was done while an intern at Microsoft Research.

http://dx.doi.org/10.1016/j.disc.2014.01.016

the inequality (1.1) was generalized to non-linear setting under the stronger assumption that every code coordinate (and not just information coordinates) has locality r.

In this paper we strengthen the results of [4,5] and establish the inequality (1.1) for general non-linear (r, d)-codes. We then use our lower-bound argument to derive structure theorems for optimal non-linear codes. Our results imply that lower bounds for parity locality of optimal (r, d)-codes that were derived in [2] in the linear setting also apply to non-linear codes. Therefore the answers to the two questions above are both negative.

1.1. Our techniques

Our new proof of the bound (1.1) follows the same high level strategy as the proofs in [2,5]. We assume that the bound (1.1) is violated and use an iterative argument to arrive at a code that violates the distance bound. Unlike [5], our iterative steps use elementary coordinate restrictions instead of entropy inequalities. This makes it easier to use our argument as a basis for structure theorems.

The main technical problem that we have to address going from the lower bound to structure theorems is that of reversibility of the local constraints. In linear codes, any local constraint on coordinates in the code must be a linear constraint, and linear constraints are trivially reversible, in that knowing all but 1 coordinate in the constraint always determines that coordinate, regardless of the identity of that 1 coordinate. However, for non-linear codes it is possible to have local constraints that are not reversible. For example, it is possible for the coordinates $\{i, i'\}$ to determine the coordinate i'', but for the coordinates $\{i', i''\}$ to not determine the coordinate i. However, we show that for optimal (r, d)-codes, even in the non-linear case, all locality constraints must be reversible. Once this is established, the structural results of [2] (and thus the parity locality lower bounds) can then be extended to the non-linear case.

2. Preliminaries

We will first fix some notation, then define the objects we will be considering.

2.1. Notation

Throughout, we consider codes which may be non-linear over an arbitrary alphabet Σ, where $|\Sigma| = q \geq 2$ is an arbitrary integer. Given two vectors $\vec{x}, \vec{y} \in \Sigma^n$, $\Delta(\vec{x}, \vec{y})$ will denote the unnormalized Hamming distance between $\vec{x}$ and $\vec{y}$. For an integer $n \geq 0$, $[n]$ denotes the set $\{1, \dots, n\}$, where $[0]$ will be understood as the empty-set. For $S \subseteq [n]$, we will denote $\vec{x}|_S$ for the sequence of symbols in $\vec{x}$ with coordinates in S. When $S = \{i\}$ we will just write $\vec{x}|_i$. For disjoint sets A and B, we write $A \sqcup B$ to denote their disjoint union.

2.2. Definitions

Recall the definition of a code, which we do not assume to be linear.

Definition 2.1. An $(n, K, d)_q$ code is a subset $\mathcal{C} \subseteq \Sigma^n$ with size $|\mathcal{C}| = K$, such that for any $\vec{x} \neq \vec{y} \in \mathcal{C}$, $\Delta(\vec{x}, \vec{y}) \geq d$. If $\mathcal{C}' \subseteq \mathcal{C}$ then $\mathcal{C}'$ is a *sub-code* of $\mathcal{C}$. The parameter n will be referred to as the *block-length*, $k = \log_q K$ the *dimension* and d the *distance*.

The code is *systematic* if $k \in \mathbb{Z}$, and there is an encoding function $\mathrm{Enc} : \Sigma^k \to \Sigma^n$ such that for $\vec{x} \in \Sigma^k$, $\mathrm{Enc}(\vec{x})|_i = \vec{x}|_i$, for $i \in [k]$.

A code is *maximum distance separable (MDS)* if $n = \log_q K + d - 1$.

A systematic code takes on all q^k values in its first k coordinates, and the values of these coordinates determine the rest of the codeword. The first k coordinates of the codewords are thus referred to as the information symbols, other coordinates will be called parity symbols. This work will be interested in codes with local constraints on the information symbols.

Definition 2.2. A systematic $(n, K, d)_q$ code has *information locality r* if for every $i \in [k]$, there is a size $\leq r$ subset $S \subseteq [n]\backslash\{i\}$ such that for any $\vec{x} \in \mathcal{C}$, $\vec{x}|_i$ is determined by $\vec{x}|_S$.

Other symbols, other than the information symbols, can also have locality, and occasionally we will use this.

3. Locality lower bounds

In this section we establish lower bounds on the block-length of codes with small information locality. We will then prove structural results for codes meeting this lower bound. As we will often use it, we now prove the Singleton bound. Recall that MDS codes are those that exactly meet this bound.